

Hand Gestures for Hearing Impaired and Inarticulate People using CNN

T. R. Shinde
M.E. in Computer Engineering
Siddhant College of Engineering
Sudumbre, Pune

Shyam Gupta, PhD
Professor in Computer Dept.
Siddhant College Engineering
Sudumbre, Pune

ABSTRACT

Gesture languages are an extremely important communication tool for many dumb and hard-of-hearing people. Sign languages are the native languages of the Deaf and Dumb community and provide full access to communication. Thanks to hearing ability we are able to perceive thoughts of every different. However, what if one fully cannot hear something and eventually cannot speak. So, the Gesture Language is that the main human action tool for hearing impaired and mute folks, associate degraded additionally to make sure a freelance life for them, the automated interpretation of linguistic communication is an in-depth analysis space. With the utilization of image process and artificial intelligence, several techniques and algorithms have been developed during this space. Each linguistic communication recognition system is trained for recognizing the gestures/signs and changing them into needed pattern. The projected system aims to produce speech to unarticulated, during this paper the double two-handed Indian linguistic communication is captured as a series of pictures and it's processed with the assistance of Python so it can produce speech and text as output.

General Terms

Convolutional Neural Network, Threshold, Sign language.

Keywords

Machine learning, Convolutional Neural Network, Hand Gesture, Threshold.

1. INTRODUCTION

Gesture languages square measure vivid on wide and world level. There square measure multiple gesture languages in world that square measure regular in use that square measure ASL (American Sign Language) ISL (Indian Sign Language), BSL (Bangla Sign Language), MSL (Malaysian Sign Language). These languages square measure engineered and developed with innumerable work and sensible testing with intention of practicability to the deaf and dumb persons. Any language is made with its word and it's that means. Sign Language is made as \Sign" and \Action of That Sign". Because here we tend to don't seem to be ready to create them perceive that means of sign by writing word. As they're deaf and might not listen from birth therefore we are able to not teach them words. Here we are using some American Sign Language (ASL) signs for prediction of hand gesture. As some used Artificial Neural Network (ANN) as a classifier for gesture recognition but only for detection of alphabets and having accuracy of 95% but in this system, we are using CNN as classifier for detection of double handed hand gestures which gives more accuracy than ANN.

2. BACKGROUND

A novel sign language learning method which do use of region of interest (ROI) segmentation pre-processing of input data through an object detection network. As the input, 2D image frames are sampled and concatenated into a wide image. The sign language learning is implemented with a convolutional neural network (CNN). 12 sign gestures are tested through a 2D camera. But it holds two limitations i.t first is an area of hand is so small that training data should be large and second is cannot add new hand gestures to be learned [1]. An original hand gesture recognition method based on DLSTM is presented. But this method is having one limitation as not worked with merged feature extracted from the hand skeleton with RGB stream information [2]. Bangla Sign Alphabets using Freeman Chain Code to extract the feature vector and Artificial Neural Network (ANN) for identification. But not recognize hand gestures in the form of words [3]. SIFT for feature extraction and Convolutional Neural Network (CNN) as a classifier for detecting Bangla Sign Languages but it cannot detect two handed gestures [4]. Study of performance for Thai sign Language identification using Kinect Sensors using various distances to find distance for accuracy. Uses fix distance 0.8 – 1.2 meter to detect gestures not more than that [5]. Compared for pre-processing, Gaussian Hidden Markov Model and Multinomial Hidden Markov Model used for edge detection. But big flaw in dataset, where signers are using gloves [6]. A novel approach to hand identification based on edge of binary images using 2D camera is introduced. They used four different techniques for feature extraction and ANN is used to classifier signs. But cannot determine words from ASL [7]. A sentence-based sign language recognition system based on motion data. System uses KNN solution and HMMs [8]. A tool which will provide communication opportunity for common people to understand mode of communication. They used CNN classifier and proposed a tool that will capture the ASL gestures made by deaf and dumb people in real time and classify those gestures into text and voice [9]. The ANN is used for training and classifies the hand sign images. The proposed method is tested using various hand sign images and results are presented to demonstrate the efficiency and effectiveness [10].

3. PROPOSED METHODOLOGY

We are offering a gadget with a view to use gadget learning set of rules i.e. CNN Convolutional Neural Network. Our planned version can be educated with round 100-500 images of hand gesture and with growing epoch as a way to increase accuracy. The reason we are the use of CNN is that it has a couple of layers as a result it will help into schooling model with smooth manner. We will use Open Computer Vision Technology concurrently to interact with digicam, to take stay input from camera. We will set and different symptoms with photos and that pictures could be trained with set of rules.

Person can have to carry out register in front of camera. After taking live enter from camera the sign might be recognized.

Recognized sign will display textual content output and it will be translated to audio sound. So, system will work as Sign to Speech. A random man or woman if visited to deaf person and if deaf character is in problem and looking to explain it then it's far very difficult to recognize what exactly he is trying to say. Delay in detecting his Sign language can turn into big critical trouble for that deaf character. This kind of people can't spend everyday life. They face communication issues at every point. Also, they get limitations and limitations to their desires and expert aims. Hence, they get demotivated and Inferiority Complex. This machine will definitely can grow to be step into innovation of this global degree hassle solution. Our device can be prototype and proof of concept for global stage solution. This machine can be used by Deaf and Dumb persons and also regular person can have this machine with them and deaf character can carry out register the front of camera and signal can be transformed to txt or speech. Objective is to provide them capability to be expressive in thoughts and thoughts. They can get helped in increasing their motivation and condense and it will help them to suppose definitely and to overcome that bodily disability. To develop gadget with the usage of trendy technologies and tools we're keeping objective to conquer from this worldwide level trouble.

3.1 Input Processing

Scanned image is taken as an snap for identification and preprocessing is performed on the scanned image input. The steps enclosed in preprocessing are.

RGB to HSV Conversion

R, G, B in RGB are all co-related to the color luminance (what we loosely call intensity), i.e., We cannot separate color information from luminance. HSV or Hue Saturation Value is used to separate image luminance from color information. This makes it easier when we are working on or need luminance of the image/frame Gray scale could be a vary of monochromatic shades from black to white. Therefore, a grey scale image contains solely reminder grey and no color.

Edge detection

For locating the picture process technique is Edge detection boundaries of objects at intervals pictures. It works by police investigation discontinuities in brightness. Edge detection is employed for image segmentation and information extraction in areas like image process, pc vision, and machine vision

Threshold

Thresholding is that the simplest technique of segmenting image. From a gray scale image, thresholding is wont to produce binary pictures. The simplest thresholding ways replace every pel in a picture with a black pel if the image intensity is a smaller amount than some fastened constant

Architecture

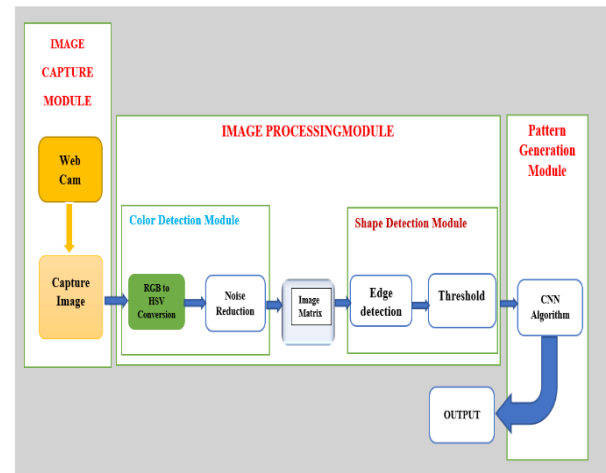


Fig. 1 System Architecture

Algorithm

In this proposed system we are using Convolutional Neural Network (CNN) algorithm as a classifier for hand gesture recognition. CNN are complex feed forward neural networks. CNNs are used for image classification and recognition because of its high accuracy. It does not contain hidden layers like ANN instead of having different filters/layers through which our image is passes.

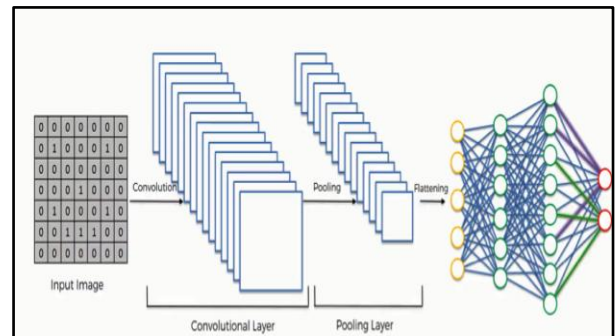


Fig.2 CNN Model

3.2 Mathematical Model

Mathematical model of the proposed system

INPUT: -

Let S is the Whole System Consist of

S- {I, P, O}

Where, I =input. I= {U, Q}

U = User U = {u1, u2,..... un}

Q = Query Q'(q1, q2,.....qn)

P= Process P = {CNN}

CNN = Convolutional Neural Network

OUTPUT: The predicted result will be the output of the system

Prediction = model. Predict (test data)

Accuracy = (accuracy score(Y test, Y red)*100)

4. EXPERIMENTAL RESULTS

In this section, experimental tests performed to evaluate performance of the proposed approach.

- CNN holds four layers:
 1. Convolutional Layer
 2. ReLu Layer
 3. Pooling Layer
 4. Fully Connected Layer

In this proposed method, the CNN algorithm is used as a classification to understand hand gesture. CNN are deep neural networks for feedback. Due to its high accuracy, CNNs are used to classify and recognize image. It does not have hidden layers, such as ANN, rather than various filters / layers that pass through the image.

Proposed System following steps -

1. To create image data set, hand gesture images are captured in threshold form to get exact hand shape.
2. 2000 images per sign hence 10,000 images for 5 gestures
3. Threshold converts image RGB to black and white color only
4. Then we get exact shape of hand
5. Then we define CNN model and add layers in it.
6. Split data into testing as 30% images and training 70% images, then train model
7. Trained model is then saved
8. The model loaded and through webcam we send input to model and model predicts it's sign on camera screen.
9. To hear the predicted text, we assigned „C” button
10. On pressing „C” we can hear the sign text

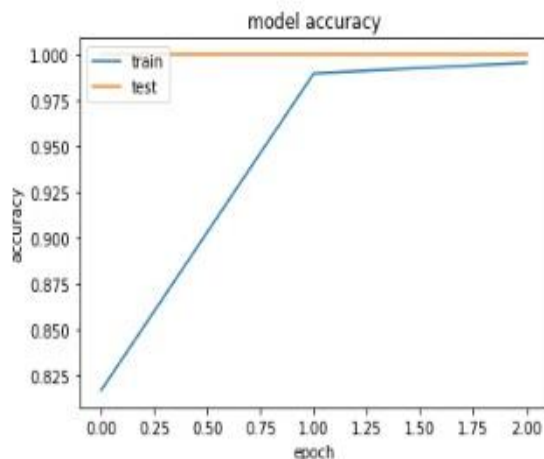


Fig. 3 Accuracy results on proposed Dataset

Below table gives accuracy using Convolutional Neural Network and Artificial Neural Network applying on dataset which contains about 10,000 images. Using ANN for detection of alphabets is 95% and using CNN for words or statement it is 98%.

Table 1. Comparison between CNN and ANN

Comparison Between CNN and ANN	
Algorithm	Accuracy
CNN (Used for detection of words)	98%
ANN (Used for detection of alphabets only)	95%

In proposed approach, in front of webcam of computer using OpenCV disable person can do the hand gestures. Shape of hand gesture will be captured by webcam and tested this data for validation. After capturing it will display its related textual message and also gives same output in audio format. Fig. 4 shows a hand gesture which gives output as “All the best”.

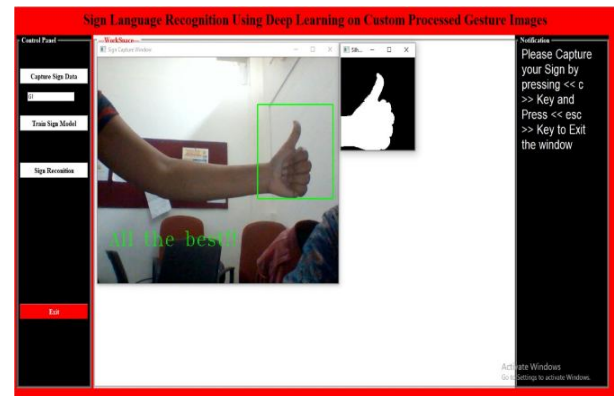


Fig. 4 Predicting Hand gesture for Indian Sign Language sentence

5. CONCLUSION

The proposed system intention to offer speech to speechless, in this paper the double surpassed Indian Sign Language is captured as a series of pictures and it is processed with the help of Python and then it's converted to speech and text. With the help of photo processing and Artificial Intelligence, many techniques and algorithms were developed in this area. Every sign language recognition device is skilled for spotting the signs and converting them into required pattern. This proposed system is only able to detect hand gesture and not whole-body gesture. There are many improvements to be made into this device. Currently it can't be able to detect whole body gesture so in future we can extend this project. Already CNN algorithm is popular algorithm in image processing system and it will definitely make huge impact on Indian Sign Language.

6. ACKNOWLEDGMENTS

Special thanks to my guide Dr. Shyam Gupta for his superior guidance, Professor, Siddhant College of Engineering, Pune

7. REFERENCES

- [1] Sunmok Kim, Yangho Ji and Ki-Beak Lee 2018 Second International Conference on Robotic Computing, An Effective Sign Language learning with object classification using ROI.
- [2] Danilo Ayola, Member IEEE, Macro Bernardi, Student

Member, IEEE, Luigi Cinque Senior Member, IEEE, Gian Luca Foresti, Senior Member, IEEE and Cristiano Massaroni, Student Member, IEEE, 2018 IEEE, Exploiting Recurrent Neural Networks and Leap Motion Controller for the Recognition of Sign Language and Semaphoric Hand Gestures.

- [3] Bangladesh Mohammad Mahadi Hasan, Md. Khaliluzzaman, Shabiba Akhter Himel and Rukhsat Tasneem Chowdhury, 4th International Conference on advances in Electrical Engineering 28-30 September, 2017, Dhaka, Hand Sign Language Recognition for Bangla Alphabet Based and ANN.
- [4] Shirin Shanta, Saif Taifur Anwar, Md. Rayhanul Kabir, IEEE 9th ICCCNT 2018 Bangla Sign Detection using SIFT and CNN.
- [5] Chana Chansri, Jakkree Srinonchat 2016 IEEE Thesis UMI order Number: GAX95-09398, University of Washington, Reliability and Accuracy of Thai Sign Language Recognition with Kinect Sensor.
- [6] Suharjito, Herman Guna wanm Narada Thiracitta, Gunawan Witjaksono, The 1st 2018 INAPR International Conference, The Comparison of some Hidden Markov

Models for Sign Language Recognition.

- [7] Ariya Thongtawee, Onamon Pinsanoh, Yuttana Kitjaidure 2018 Biomedical Engineering International Conference (BMEiCON-2018) A Novel Feature Extraction for American Sign Language Recognition
- [8] Mohamed Hassan, Khaled Assaleh, Tamer Shanableh, 2016 International conference on Computational Science and Computational Intelligence, User-dependent Sign Language Recognition Using Motion Detection Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.
- [9] Prateek SG, Siddarth R, P.G.Sunitha Hiremath, Jagdeesh J, Smitha Y, Neha Tarannum Pendari, International Conference on Advances in Computing, Communication Control and Networking (ICACCCN2018), @2018 IEEE, Dynamic Tool for American Sign Language Finger Spelling Interpreter.
- [10] Prateek SG Jagadeesh I International Conference on Advances in Computing, Communication Control and Networking (ICACCCN2018), Dynamic Tool for American Sign Language